

RECEIVED

MAR - 7 2002

TECH CENTER 1600/2900



1646

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/825,751A

TIME: 11:30:33

Input Set : A:\Cura-251.app

Output Set: N:\CRF3\02212002\I825751A.raw

P.S.

ENTERED

3 <110> APPLICANT: CuraGen Corporation
 4 Vernet, Corine A.M.
 5 Fernandes, Elma R
 6 Taupier, Raymond J
 7 Quinn, Kerry E
 8 Spytek, Kimberly A
 9 Rastelli, Luca
 10 Herrman, John L
 12 <120> TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
 14 <130> FILE REFERENCE: 15966-750
 16 <140> CURRENT APPLICATION NUMBER: 09/825,751A
 C--> 17 <141> CURRENT FILING DATE: 2001-04-30
 19 <150> PRIOR APPLICATION NUMBER: 60/194,314
 20 <151> PRIOR FILING DATE: 2000-04-03
 22 <150> PRIOR APPLICATION NUMBER: 60/225,693
 23 <151> PRIOR FILING DATE: 2000-08-16
 25 <160> NUMBER OF SEQ ID NOS: 85
 27 <170> SOFTWARE: PatentIn Ver. 2.1
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 1852
 31 <212> TYPE: DNA
 32 <213> ORGANISM: Homo sapiens
 34 <220> FEATURE:
 35 <221> NAME/KEY: CDS
 36 <222> LOCATION: (208)..(1698)
 38 <400> SEQUENCE: 1
 39 cggatgactc ccgagaaggt gagccctca cccacatgct aagagcccct tctgggccac 60
 41 ccagatccat ctccgcactg cctgggtctc tgagtctcag gctccccctg agagcctggg 120
 43 tggccctgga ccctgccagc ctggggcctg ggcttttctc ccctgggggc cttgagtgtg 180
 45 gccagggctc tggcgattgt gtggtga cag aag cca tgt ctg caa cgc ctg cca 234
 46 Gln Lys Pro Cys Leu Gln Arg Leu Pro
 47 1 5
 49 tcc gca gac gtg aat gag tgt gca gag aac cct ggc gtc tgc act aac 282
 50 Ser Ala Asp Val Asn Glu Cys Ala Glu Asn Pro Gly Val Cys Thr Asn
 51 10 15 20 25
 53 ggc gtc tgt gtc aac acc gat gga tcc ttc cgc tgt gag tgt ccc ttt 330
 54 Gly Val Cys Val Asn Thr Asp Gly Ser Phe Arg Cys Glu Cys Pro Phe
 55 30 35 40
 57 ggc tac agc ctg gac ttc act ggc atc aac tgt gtg gac aca gac gag 378
 58 Gly Tyr Ser Leu Asp Phe Thr Gly Ile Asn Cys Val Asp Thr Asp Glu
 59 45 50 55
 61 tgc tct gtc ggc cac ccc tgt ggg caa ggg aca tgc acc aat gtc atc 426
 62 Cys Ser Val Gly His Pro Cys Gly Gln Gly Thr Cys Thr Asn Val Ile

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/825,751A

TIME: 11:30:33

Input Set : A:\Cura-251.app

Output Set: N:\CRF3\02212002\I825751A.raw

63	60	65	70	
65	gga ggc ttc gaa tgt gcc tgt gct gac ggc ttt gag cct ggc ctc atg	474		
66	Gly Gly Phe Glu Cys Ala Cys Ala Asp Gly Phe Glu Pro Gly Leu Met			
67	75	80	85	
69	atg acc tgc gag gac atc gac gaa tgc tcc ctg aac ccg ctg ctc tgt	522		
70	Met Thr Cys Glu Asp Ile Asp Glu Cys Ser Leu Asn Pro Leu Leu Cys			
71	90	95	100	105
73	gcc ttc cgc tgc cac aat acc gag ggc tcc tac ctg tgc acc tgt cca	570		
74	Ala Phe Arg Cys His Asn Thr Glu Gly Ser Tyr Leu Cys Thr Cys Pro			
75	110	115	120	
77	gcc ggc tac acc ctg cgg gag gac ggg gcc atg tgt cga gat gtg gac	618		
78	Ala Gly Tyr Thr Leu Arg Glu Asp Gly Ala Met Cys Arg Asp Val Asp			
79	125	130	135	
81	gag tgt gca gat ggt cag cag gac tgc cac gcc cgg ggc atg gag tgc	666		
82	Glu Cys Ala Asp Gly Gln Gln Asp Cys His Ala Arg Gly Met Glu Cys			
83	140	145	150	
85	aag aac ctc atc ggt acc ttc gcg tgc gtc tgt ccc cca ggc atg cgg	714		
86	Lys Asn Leu Ile Gly Thr Phe Ala Cys Val Cys Pro Pro Gly Met Arg			
87	155	160	165	
89	ccc ctg cct ggc tct ggg gag ggc tgc aca gat gac aat gaa tgc cac	762		
90	Pro Leu Pro Gly Ser Gly Glu Gly Cys Thr Asp Asp Asn Glu Cys His			
91	170	175	180	185
93	gct cag cct gac ctc tgt gtc aac ggc cgc tgt gtc aac acc gcg ggc	810		
94	Ala Gln Pro Asp Leu Cys Val Asn Gly Arg Cys Val Asn Thr Ala Gly			
95	190	195	200	
97	agc ttc cgg tgc gac tgt gat gag gga ttc cag ccc agc ccc acc ctt	858		
98	Ser Phe Arg Cys Asp Cys Asp Glu Gly Phe Gln Pro Ser Pro Thr Leu			
99	205	210	215	
101	acc gag tgc cac gac atc cgg cag ggg ccc tgc ttt gcc gag gtg ctg	906		
102	Thr Glu Cys His Asp Ile Arg Gln Gly Pro Cys Phe Ala Glu Val Leu			
103	220	225	230	
105	cag acc atg tgc cgg tct ctg tcc agc agc agt gag gct gtc acc agg	954		
106	Gln Thr Met Cys Arg Ser Leu Ser Ser Ser Ser Glu Ala Val Thr Arg			
107	235	240	245	
109	gcc gag tgc tgc tgt ggg ggt ggc cgg ggc tgg ggg ccc cgc tgc gag	1002		
110	Ala Glu Cys Cys Cys Gly Gly Gly Arg Gly Trp Gly Pro Arg Cys Glu			
111	250	255	260	265
113	ctc tgt ccc ctg ccc ggc acc tct gcc tac agg aag ctg tgc ccc cat	1050		
114	Leu Cys Pro Leu Pro Gly Thr Ser Ala Tyr Arg Lys Leu Cys Pro His			
115	270	275	280	
117	ggc tca ggc tac act gct gag ggc cga gat gta gat gaa tgc cgt atg	1098		
118	Gly Ser Gly Tyr Thr Ala Glu Gly Arg Asp Val Asp Glu Cys Arg Met			
119	285	290	295	
121	ctt gct cac ctg tgt gct cat ggg gag tgc atc aac agc ctt ggc tcc	1146		
122	Leu Ala His Leu Cys Ala His Gly Glu Cys Ile Asn Ser Leu Gly Ser			
123	300	305	310	
125	ttc cgc tgc cac tgt cag gcc ggg tac aca ccg gat gct act gct act	1194		
126	Phe Arg Cys His Cys Gln Ala Gly Tyr Thr Pro Asp Ala Thr Ala Thr			
127	315	320	325	

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/825,751A

TIME: 11:30:33

Input Set : A:\Cura-251.app

Output Set: N:\CRF3\02212002\I825751A.raw

```

129 acc tgc ctg gat atg gat gag tgc agc cag gtc ccc aag cca tgt acc 1242
130 Thr Cys Leu Asp Met Asp Glu Cys Ser Gln Val Pro Lys Pro Cys Thr
131 330 335 340 345
133 ttc ctc tgc aaa aac acg aag ggc agt ttc ctg tgc agc tgt ccc cga 1290
134 Phe Leu Cys Lys Asn Thr Lys Gly Ser Phe Leu Cys Ser Cys Pro Arg
135 350 355 360
137 ggc tac ctg ctg gag gag gat ggc agg acc tgc aaa gac ctg gac gaa 1338
138 Gly Tyr Leu Glu Glu Asp Gly Arg Thr Cys Lys Asp Leu Asp Glu
139 365 370 375
141 tgc acc tcc cgg cag cac aac tgt cag ttc ctc tgt gtc aac act gtg 1386
142 Cys Thr Ser Arg Gln His Asn Cys Gln Phe Leu Cys Val Asn Thr Val
143 380 385 390
145 ggc gcc ttc acc tgc cgc tgt cca ccc ggc ttc acc cag cac cac cag 1434
146 Gly Ala Phe Thr Cys Arg Cys Pro Pro Gly Phe Thr Gln His His Gln
147 395 400 405
149 gcc tgc ttc gac aat gat gag tgc tca gcc cag cct ggc cca tgt ggt 1482
150 Ala Cys Phe Asp Asn Asp Glu Cys Ser Ala Gln Pro Gly Pro Cys Gly
151 410 415 420 425
153 gcc cac ggg cac tgc cac aac acc ccg ggc agc ttc cgc tgt gaa tgc 1530
154 Ala His Gly His Cys His Asn Thr Pro Gly Ser Phe Arg Cys Glu Cys
155 430 435 440
157 cac caa ggc ttc acc ctg gtc agc tca ggc cat ggc tgt gaa gat gtg 1578
158 His Gln Gly Phe Thr Leu Val Ser Ser Gly His Gly Cys Glu Asp Val
159 445 450 455
161 aat gaa tgt gat ggg ccc cac cgc tgc cag cat ggc tgt cag aac cag 1626
162 Asn Glu Cys Asp Gly Pro His Arg Cys Gln His Gly Cys Gln Asn Gln
163 460 465 470
165 cta ggg ggc tac cgc tgc agc tgc ccc cag ggt ttc acc cag cac tcc 1674
166 Leu Gly Gly Tyr Arg Cys Ser Cys Pro Gln Gly Phe Thr Gln His Ser
167 475 480 485
169 cag tgg gcc cag tgt gtg ggt gag tgaaaaagggc tgggaagaag ctgggccctc 1728
170 Gln Trp Ala Gln Cys Val Gly Glu
171 490 495
173 caccagaatc tgctcagagc aggcgactaa cagacgccac cctgcaagat gatgtgacaa 1788
175 gcacaattat ctaaagattg aacaggccag cccagaagat gagaatgagt gtgccctgtc 1848
177 gccc 1852
180 <210> SEQ ID NO: 2
181 <211> LENGTH: 497
182 <212> TYPE: PRT
183 <213> ORGANISM: Homo sapiens
185 <400> SEQUENCE: 2
186 Gln Lys Pro Cys Leu Gln Arg Leu Pro Ser Ala Asp Val Asn Glu Cys
187 1 5 10 15
189 Ala Glu Asn Pro Gly Val Cys Thr Asn Gly Val Cys Val Asn Thr Asp
190 20 25 30
192 Gly Ser Phe Arg Cys Glu Cys Pro Phe Gly Tyr Ser Leu Asp Phe Thr
193 35 40 45
195 Gly Ile Asn Cys Val Asp Thr Asp Glu Cys Ser Val Gly His Pro Cys
196 50 55 60

```

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/825,751A

TIME: 11:30:33

Input Set : A:\Cura-251.app

Output Set: N:\CRF3\02212002\I825751A.raw

```

198 Gly Gln Gly Thr Cys Thr Asn Val Ile Gly Gly Phe Glu Cys Ala Cys
199 65 70 75 80
201 Ala Asp Gly Phe Glu Pro Gly Leu Met Met Thr Cys Glu Asp Ile Asp
202 85 90 95
204 Glu Cys Ser Leu Asn Pro Leu Leu Cys Ala Phe Arg Cys His Asn Thr
205 100 105 110
207 Glu Gly Ser Tyr Leu Cys Thr Cys Pro Ala Gly Tyr Thr Leu Arg Glu
208 115 120 125
210 Asp Gly Ala Met Cys Arg Asp Val Asp Glu Cys Ala Asp Gly Gln Gln
211 130 135 140
213 Asp Cys His Ala Arg Gly Met Glu Cys Lys Asn Leu Ile Gly Thr Phe
214 145 150 155 160
216 Ala Cys Val Cys Pro Pro Gly Met Arg Pro Leu Pro Gly Ser Gly Glu
217 165 170 175
219 Gly Cys Thr Asp Asp Asn Glu Cys His Ala Gln Pro Asp Leu Cys Val
220 180 185 190
222 Asn Gly Arg Cys Val Asn Thr Ala Gly Ser Phe Arg Cys Asp Cys Asp
223 195 200 205
225 Glu Gly Phe Gln Pro Ser Pro Thr Leu Thr Glu Cys His Asp Ile Arg
226 210 215 220
228 Gln Gly Pro Cys Phe Ala Glu Val Leu Gln Thr Met Cys Arg Ser Leu
229 225 230 235 240
231 Ser Ser Ser Ser Glu Ala Val Thr Arg Ala Glu Cys Cys Cys Gly Gly
232 245 250 255
234 Gly Arg Gly Trp Gly Pro Arg Cys Glu Leu Cys Pro Leu Pro Gly Thr
235 260 265 270
237 Ser Ala Tyr Arg Lys Leu Cys Pro His Gly Ser Gly Tyr Thr Ala Glu
238 275 280 285
240 Gly Arg Asp Val Asp Glu Cys Arg Met Leu Ala His Leu Cys Ala His
241 290 295 300
243 Gly Glu Cys Ile Asn Ser Leu Gly Ser Phe Arg Cys His Cys Gln Ala
244 305 310 315 320
246 Gly Tyr Thr Pro Asp Ala Thr Ala Thr Thr Cys Leu Asp Met Asp Glu
247 325 330 335
249 Cys Ser Gln Val Pro Lys Pro Cys Thr Phe Leu Cys Lys Asn Thr Lys
250 340 345 350
252 Gly Ser Phe Leu Cys Ser Cys Pro Arg Gly Tyr Leu Leu Glu Glu Asp
253 355 360 365
255 Gly Arg Thr Cys Lys Asp Leu Asp Glu Cys Thr Ser Arg Gln His Asn
256 370 375 380
258 Cys Gln Phe Leu Cys Val Asn Thr Val Gly Ala Phe Thr Cys Arg Cys
259 385 390 395 400
261 Pro Pro Gly Phe Thr Gln His His Gln Ala Cys Phe Asp Asn Asp Glu
262 405 410 415
264 Cys Ser Ala Gln Pro Gly Pro Cys Gly Ala His Gly His Cys His Asn
265 420 425 430
267 Thr Pro Gly Ser Phe Arg Cys Glu Cys His Gln Gly Phe Thr Leu Val
268 435 440 445
270 Ser Ser Gly His Gly Cys Glu Asp Val Asn Glu Cys Asp Gly Pro His

```

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/825,751A

TIME: 11:30:33

Input Set : A:\Cura-251.app

Output Set: N:\CRF3\02212002\I825751A.raw

```

271      450      455      460
273 Arg Cys Gln His Gly Cys Gln Asn Gln Leu Gly Gly Tyr Arg Cys Ser
274 465      470      475      480
276 Cys Pro Gln Gly Phe Thr Gln His Ser Gln Trp Ala Gln Cys Val Gly
277      485      490      495
279 Glu
284 <210> SEQ ID NO: 3
285 <211> LENGTH: 379
286 <212> TYPE: DNA
287 <213> ORGANISM: Homo sapiens
289 <220> FEATURE:
290 <221> NAME/KEY: CDS
291 <222> LOCATION: (1)..(378)
293 <400> SEQUENCE: 3
294 gga ggg cct gtg att cta ctg cag gca ggc acc ccc cac aac ctc aca 48
295 Gly Gly Pro Val Ile Leu Leu Gln Ala Gly Thr Pro His Asn Leu Thr
296 1 5 10 15
298 tgc cgg gcc ttc aat gcg aag cct gct gcc acc atc atc tgg ttc cgg 96
299 Cys Arg Ala Phe Asn Ala Lys Pro Ala Ala Thr Ile Ile Trp Phe Arg
300 20 25 30
302 gac ggg acg cag cag gag ggc gct gtg gcc agc acg gaa ttg ctg aag 144
303 Asp Gly Thr Gln Gln Glu Gly Ala Val Ala Ser Thr Glu Leu Leu Lys
304 35 40 45
306 gat ggg aag agg gag acc acc gtg agc caa ctg ctt att aac ccc acg 192
307 Asp Gly Lys Arg Glu Thr Thr Val Ser Gln Leu Leu Ile Asn Pro Thr
308 50 55 60
310 gac ctg gac ata ggg cgt gtc ttc act tgc cga agc atg aac gaa gcc 240
311 Asp Leu Asp Ile Gly Arg Val Phe Thr Cys Arg Ser Met Asn Glu Ala
312 65 70 75 80
314 atc cct agt ggc aag gag act tcc atc gag ctg gat gtg cac cac cct 288
315 Ile Pro Ser Gly Lys Glu Thr Ser Ile Glu Leu Asp Val His His Pro
316 85 90 95
318 cct aca gtg acc ctg tcc att gag cca cag acg ggg cag gag ggt gag 336
319 Pro Thr Val Thr Leu Ser Ile Glu Pro Gln Thr Gly Gln Glu Gly Glu
320 100 105 110
322 cgt gtt gtc ttt acc tgc cag gcc aca gcc aac ccc gag atc t 379
323 Arg Val Val Phe Thr Cys Gln Ala Thr Ala Asn Pro Glu Ile
324 115 120 125
327 <210> SEQ ID NO: 4
328 <211> LENGTH: 126
329 <212> TYPE: PRT
330 <213> ORGANISM: Homo sapiens
332 <400> SEQUENCE: 4
333 Gly Gly Pro Val Ile Leu Leu Gln Ala Gly Thr Pro His Asn Leu Thr
334 1 5 10 15
336 Cys Arg Ala Phe Asn Ala Lys Pro Ala Ala Thr Ile Ile Trp Phe Arg
337 20 25 30
339 Asp Gly Thr Gln Gln Glu Gly Ala Val Ala Ser Thr Glu Leu Leu Lys
340 35 40 45

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 02/21/2002

PATENT APPLICATION: US/09/825,751A

TIME: 11:30:34

Input Set : A:\Cura-251.app

Output Set: N:\CRF3\02212002\I825751A.raw

L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:2236 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57

L:2237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57

L:2291 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61